Molecular dynamics simulations in rational protein design
Creveld, Lucia Deborah

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2001

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
CHAPTER THREE
IDENTIFICATION OF FUNCTIONAL AND UNFOLDING MOTIONS OF CUTINASE AS OBTAINED FROM MOLECULAR DYNAMICS COMPUTER SIMULATIONS

ABSTRACT

The implementation of cutinase from *Fusarium solani pisi* as a fat-stain removing ingredient in laundry washing is hampered by its unfolding in the presence of anionic surfactants. In this work we present molecular dynamics (MD) computer simulations of cutinase and analysis procedures to distinguish the movements related to its functional behavior (e.g. substrate binding) from those related to the unfolding of the enzyme. Two kinds of MD-simulations were performed: a simulation mimicking the thermal motion at room temperature, and several simulations in which unfolding is induced either by high temperature or by using a modified water-protein interaction potential. Essential dynamics analyses [A. Amadei et al. Proteins 17: 412-425, 1993] on the simulations identify distinct regions in the molecular structure of cutinase in which the motions occur for function and initial unfolding. The unfolding in various simulations starts in a similar way, suggesting that mutations in the regions involved might stabilize the enzyme without affecting its functionality.

Chapter 3

3.1. INTRODUCTION

Cutinases are enzymes with a molecular mass of 22-25 kDa that supposedly hydrolyze ester bonds of the cutin polymer in the plant cell wall.\(^1\) X-ray diffraction studies of the *Fusarium solani pisi* enzyme revealed that it adopts a parallel $\alpha/\beta$ fold and that has a catalytic triad consisting of Ser120, His188 and Asp175.\(^2\) The structure of the enzyme is shown schematically in figure 3.1. The implementation of the enzyme as a fat-stain removing ingredient in laundry washing is hampered by its unfolding in the presence of anionic surfactants.

The hypothesis underlying this study is that the apolar tails of the surfactants are capable of penetrating the protein at weak spots on its surface, in particular at hydrophobic patches which are transiently formed during the motion of the protein. In our approach, insight into the internal mobility of the protein on an atomic level is obtained from molecular dynamics (MD) computer simulations.\(^3\) As it is not straightforward to simulate the enzyme in the presence of anionic surfactants, the motion of the enzyme is simulated either at high temperature or using a modified water-protein interaction potential. We hypothesize that the weak spots found in these simulations comprise the ones responsible for the sensitivity to anionic surfactants. In this study we do not have the ambition to elucidate the complete unfolding pathway of the protein, but more modestly to find the regions in which unfolding is initiated. Knowledge of these regions can subsequently be used as one of the guidelines to design fully functional variants with improved stability or to design more efficient enzymes with the same stability.

3.2. METHODS

All calculations were carried out using the GROMOS87 simulation package.\(^4\) *Insight II* (MSI, San Diego) and *WhatIf*\(^5\) were used for visualization of snapshots and GROMOS trajectories, respectively. Secondary structure elements were assigned by the DSSP-program package.\(^6\) Solvent accessible surface area were calculated by NACCESS.\(^7\)

3.2.1. Molecular Dynamics simulations (MD)

The crystal structure at a resolution of 1.2 Å of *Fusarium solani pisi* cutinase,\(^2\) entry 1cus in the Brookhaven Protein Databank,\(^8\) served as starting structure for the simulations after the side chain amide-groups of Asn47, Asn152 and Gln71 were rotated by 180 degrees to form energetically more favorable hydrogen bonds.\(^5\) The protein and its crystal water molecules were placed in a periodic truncated octahedron of equilibrated SPC/E water.\(^9\) The minimum distance between the protein and the edge of the box was 10.0 Å. All water molecules with their oxygen closer than 2.3 Å to any non-hydrogen protein atom or crystal water oxygen were removed. In order to relieve unfavorable non-bonded interactions with the added water molecules, energy minimization was performed while the protein and the crystal water oxygen atoms were harmonically restrained with respect to the X-ray structure using a force constant of 4000 kJ mol\(^{-1}\) nm\(^{-2}\). For all MD simulations initial velocities were taken from a Maxwel-
Functional and unfolding motions of cutinase

Figure 3.1: Structure of Fusarium solani pisi cutinase. β-strands are indicated as light grey arrows in the direction of the sequence (S1 = 33-40, S2 = 67-73, S3 = 112-119, S4 = 141-148 and S5 = 166-171). α-helices are indicated as dark grey cylinders (H1 = 22-25, H2 = 28-30, H3 = 51-63, H4 = 80-85, H5 = 91-108, H6 = 121-132, H7 = 134-138, H8 = 152-155, H9 = 163-166, H10 = 175-179, H11 = 185-189, H12 = 191-196, H13 = 198-211). The active site residues: Ser120, Asp175 and His188 are indicated as cpk-model. The Cα-trace of turns and random coil regions are indicated as a ribbon.

Lian distribution at 293 K. The systems were equilibrated by harmonically restraining the protein atoms to their crystallographic positions with a force constant of 4000kJmol⁻¹nm⁻². and by applying temperature and pressure coupling¹⁰ with coupling time constants of 0.01 ps and 0.05 ps respectively. Over a period of 25 ps the harmonic restraints as well as the temperature and pressure coupling were released. The simulations were continued at constant volume. Covalent bond lengths were constrained using the SHAKE procedure¹¹ with a relative tolerance of 10⁻⁴. The time step used in the leap-frog integration scheme was 2 fs. The temperatures of the protein and
the solvent were separately maintained constant by weak coupling to a heat bath with a coupling constant of 0.10 ps. Non-bonded interactions were handled by means of a twin-range method, using short and long range cut-off radii of 8.0 Å and 12.0 Å, respectively, and a pairlist updated every 20 fs. During the simulations, a structure was saved for further analysis every 0.1 ps.

A “functional simulation” (referred to as funct-sim) was performed using the 37C4 united atom force field of GROMOS87 with the following modifications: (1) For the interaction between a carbon atom and SPC/E water oxygen the value of the C12 parameter of 421.0 was replaced by 793.3 (kcal mol$^{-1}$ Å$^{12}$)$^{1/2}$. (2) The aromatic hydrogen atoms were included explicitly with C/H atomic partial charges of $-0.14e/+0.14e$. The simulated system comprised 1839 protein atoms and 5941 water molecules. The simulation was stopped after 1505 ps (including equilibration).

A simulation at high temperature (referred to as T-unf-sim) was started from the structure at 255 ps of funct-sim, after the structure was equilibrated. During 50 ps the temperature was raised gradually by 100 K after which the simulation was continued for 1170 ps at 393 K, giving a total simulation time of 1475 ps.

Several simulations with a modified protein-solvent interaction were performed using the unmodified 37C4 force field of GROMOS87, i.e., without explicit aromatic hydrogen atoms and with the original value of the C12 parameter. The latter resulted in an unrealistically small repulsion between the oxygen atoms of water and the carbon atoms of the protein. In this paper, only one of these simulations is discussed in more detail (referred to as S-unf-sim), as all simulations show similar results. Unfortunately the first 400 ps of S-unf-sim were lost due to a disk crash. Nevertheless it was chosen as it is the only one that displays all relevant motions in a single simulation. The simulated system consisted of 1778 protein atoms and 5806 water molecules. The simulation was stopped at 1300 ps (including equilibration).

### 3.2.2. Analyses

In all analyses that require translational-rotational fitting of the structures, only Cα-carbons of the stable secondary structure elements of funct-sim were used (residues: 33–40 (strand 1), 51–63 (helix 3), 67–73 (strand 2), 91–108 (helix 5), 112–119 (strand 3), 121–132 (helix 6), 134–138 (helix 7), 141–148 (strand 4), 166–171 (helix 13)).

To assess the stability of the simulations, the root-mean-square positional deviation (RMSD) of Cα-atoms with respect to the X-ray structure, the solvent accessible surface area, and the presence of secondary structure elements were monitored. In order to compare funct-sim with the mobility in the crystal, positional root-mean-square fluctuations (RMSF) were calculated (1) for specified atoms k from the simulation as $\text{RMSF}_k = \langle (r_k - \langle r_k \rangle)^2 \rangle^{1/2}$, with $\langle r_k \rangle$ being the average position over time and $r_k$ the position at every time step and (2) for the crystal as $\text{RMSF}_k = (3B_k / 8\pi^2)^{1/2}$, with $B_k$ being the B-factor for residue k. Crystal packing contacts for residues in the X-ray-struct-
Functional and unfolding motions of cutinase

ture were assigned to residues which contain at least one atom at a distance less than 5 Å to at least one protein atom of a neighboring molecule in the crystal. The program QUILT\textsuperscript{16,17} was applied to identify the distribution of hydrophobic patches. A patch is defined here as a contiguous piece of solvent accessible surface made up of a cluster of neighboring atoms that are apolar. Some C-atoms next to =O or -OH are incorporated at the edge of patches by QUILT, which may give slightly too large patches, but does not change the distribution of the patches. H-bonds were analyzed with the routine proahb of GROMOS87 package\textsuperscript{4} using default parameters (distance hydrogen-acceptor lower than 2.5 Å, angle donor-hydrogen-acceptor greater than 135°).

The essential dynamics analysis method\textsuperscript{18} distinguishes large concerted internal motions from small random internal motions. In short the method comprises: (1) performing a translational/rotational fit of all frames in a chosen time window to a common reference, (2) calculation of the C\textalpha positional fluctuation correlation matrix: $c_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle$, where $x_i$ and $x_j$ represent all cartesian coordinates of the C\textalpha- atoms, (3) diagonalization of this matrix, and (4) ordering of eigenvectors by decreasing eigenvalue. Only motions along the first few eigenvectors describe significant motions in a protein.\textsuperscript{18-22} Further eigenvectors contain the less interesting, “near-constrained” motions. Often, the latter correspond to small Gaussian-distributed random fluctuations. In this paper we call “essential motions” these motions that are characterized by a correlation coefficient lower than 0.90 between the actually sampled distribution of a vector with the Gaussian distribution calculated from its eigenvalue. It turned out that this criterion leaves a similar amount of residual, near-constrained motion for funct-sim and S-unf-sim after extracting the large essential motions. To allow 3D-visualisation of the essential motions of funct-sim the third eigenvector was included in further analyses, even though with its correlation coefficient of 0.89 it should not. The information on the size and anisotropy of the essential motions for the C\textalpha-atoms was visualized by ellipsoids. To calculate the ellipsoids, first a 3N x 3N correlation matrix was constructed by back transformation using only the eigenvalues and -vectors of the essential motions. For each C\textalpha-atom a 3 x 3 submatrix was extracted and diagonalized. The eigenvalues and -vectors from this submatrix span the ellipsoid for this atom. The ellipsoids were visualized using an application for the program Date Explorer\textsuperscript{TM} IBM. The essential RMSF of a C\textalpha-atom is the square root of the trace of its submatrix.

3.3. RESULTS

3.3.1. Stability of molecular dynamics simulations

For all simulations the positional root-mean-square deviation (RMSD) from the X-ray structure (figure 3.2a), solvent accessible surface area (figure 3.2b-d) and secondary structure (figure 3.3) were monitored to assess the stability of the simulations. The RMSD of funct-sim increases quickly during 25 ps to 1.2 Å and from there on slowly (0.3 Å ns\textsuperscript{-1}), indicating the release of strain from the starting structure during the first 25 ps (figure 3.2a). Main chain, polar and non-polar side chain solvent accessible sur-
face areas are constant throughout the simulation (1270 ± 8, 2775 ± 15 and 4800 ± 15 respectively, figure 3.2b-d), also indicating that the simulation is stable. The main secondary structure elements (helices 3, 5, 6, 7, and 13; all strands) are well preserved in funct-sim (figure 3.3a). Smaller helical structures (helices 1, 2, 4, 8, 9, 10, 11 and 12) unfold and refold several times during the simulation.

Figure 3.2: The root mean square deviations (RMSD) of Cα-atoms (a), and accessible surface area of main chain (b) polar side chains (c), and non polar side chains (d) for funct-sim (dotted line), S-unf-sim (full line) and T-unf-sim (dashed line).

The root-mean-square fluctuation (RMSF) of funct-sim is compared to the RMSF calculated from the crystallographic B-factors of the X-ray structure in figure 3.4. Overall, the regions with high or low mobility are distributed similarly in the B-factors and in the simulation. For both, high mobility is found at one side of the binding site (179-187), and in loops (e.g. residues 27-30, 109-111) and low mobility is found for main secondary structure elements (e.g. strand 3, helix 5). The average RMSF over all residues is roughly the same, but the RMSF of the simulation has more extreme variations than that of the B-factors. The higher fluctuations of the X-ray-data for most residues can be due to static disorder and to oscillations of the molecule as a whole around its lattice position in the crystal. The higher RMSF of the MD data for some regions, e.g. at the N- and C-termini and at residues 42-45, 131, 132 and 193, in the protein can be explained from the absence of crystal contacts in the simulation, which results in a larger
Figure 3.3a: Secondary structure as a function of time in funct-sim, as assigned by the DSSP-program package. α-helices are colored black and β-strands dark grey.

conformational freedom of the protein. Not only the residues with crystal contact themselves are restrained in crystal but also, indirectly, their neighbours. The removal of crystal contacts upon using the crystal structure for a simulation may lead to a conformational adaptation towards a slightly different, more “solvent-like” structure. The largest conformational adaptations are seen during the first 255 ps of the simulation. The conformational changes for residues 33, 64-66, 152-158 (helix 8), and 172-174 even occur until 700 ps. The higher positional fluctuations of these residues in the simulation do not reveal thermal motion, but rather an interconversion between two or more conformations upon relaxation of the crystal contacts. The conformational change of residue 33 together with 64-66 is not due to the loss of the H-bond between the pi-system of Trp69 and the HN of Ala32, which was observed by NMR and, with hindsight,
also in the crystal structure. In the simulation this H-bond is very well maintained, probably because the forces are overestimated in the GROMOS forcefield. This H-bond may prevent the rearrangement of residues 27-30, the conformation of which is ill-defined by the electron density (C. Cambillau, personal communication). The conformational change for residues 152-158 involves the formation of three extra H-bonds in the main chain (data not shown) giving rise to helix 8 (152-155), which is absent in the crystal structure, but present according to the NMR data. It is gratifying to observe that this helix is spontaneously formed in the MD simulation, even though it is not stable (figure 3.3a). The conformational change of residues 172-174 is constrained by the SS-bridge of residues 171 and 179.

The RMSD of S-unf-sim (figure 3.2a) increases significantly in the course of the simulation. The accessible surface area (figure 3.2b-d) shows an increase in time, for the non-polar side chains and to a lesser extent, for the main chain atoms. Visual inspection indicates that the large motions observed before in funct-sim, namely in the binding site, and at the N- and C-termini, are present in S-unf-sim as well (figure 3.5). The rearrangements of residues 52-66, 134-141, and 151-162 are however more pronounced.
and different from the one observed in funct-sim. Some of the secondary structure of cutinase (figure 3.3b) is slowly lost in the course of S-unf-sim. Helices 4, 10 and 11, the end of helix 6 and the first part of strand 4 are lost already at 400 ps, followed by the complete unfolding of helix 1 and strand 5 at 800 ps, and of helix 7 at 1200 ps. The other unfolding simulations with modified protein-solvent interaction show a similar behavior, although secondary structure elements are lost earlier or later in time and in different order (data not shown).

The RMSD of T-unf-sim (figure 3.2a) increases smoothly until 600 ps and from there on slowly (0.3 Å ns⁻¹). Although the RMSD increases, the accessible surface area (figure 3.2b-d) does not change over time. The radius of gyration shows a similar effect: it
is constant for funct-sim and this simulation, while it increases for S-unf-sim (not shown). Visual inspection indicates that the same residues change conformation as in S-unf-sim. The N-terminal (residues 17-21) however shows a larger displacement when compared to S-unf-sim. The secondary structures in T-unf-sim show large fluctuations due to the higher temperature (figure 3.3c). Helix 1, 4 and the C-terminal part of helix 6 are lost early however, followed by parts of strands 2, 4 and 5. Helix 7 unfolds and refolds several times at the end of the simulation. Although the unfolding has a different character than in S-unf-sim and does not continue after 600 ps, the unfolding of the enzyme starts in the same regions.

Further analysis is presented for S-unf-sim only, as this simulation is most representative of the unfolding behavior by anionic surfactants (see discussion).
Figure 3.5: Two snapshots of S-unf-sim depicted in a ribbon presentation. The starting structure is colored white with green and the 600 ps structure yellow with red. Functional behavior is indicated for residues 80-90 and 179-187, and initial unfolding (colored green and red) for residues 52-66, 133-141, 151-166 and 191-213. The active site residues are indicated as orange sticks and the cysteine-bridges as yellow sticks.
3.3.2. Essential dynamics analyses

Essential dynamics analysis was performed on funct-sim for 255-1505 ps and on S-unf-sim for 400-1300 ps. In funct-sim the first three eigenvectors are essential, as defined in the methods section, and in S-unf-sim vectors 1, 2, 3, 4 and 7 are. The remaining “near-constrained” motion (RMSF) is comparable for both simulations: 0.29 Å on average per residue for funct-sim and 0.24 Å for S-unf-sim. This is minor compared with their total RMSF of 0.72 Å and 1.43 Å on average per residue, respectively (figure 3.6). Obviously, S-unf-sim contains many more essential motion, which can be ascribed to the unfolding of cutinase. The RMSF over the essential vectors of funct-sim is shown together with the RMSF of the simulation itself in figure 3.4. They differ constantly by 0.25 Å, indicating that the “near-constrained” motion is distributed proportionally over all Cα-atoms. Thus, the essential dynamics analysis indeed removed the “random” motions, leaving the most important correlated motions.

![Figure 3.6: The essential (grey) and “near constrained” (black) average root mean square fluctuation (RMSF) per Cα-atom of funct-sim and S-unf-sim.](image)

In figure 3.7 the combined essential eigenvectors of funct-sim are displayed as the ellipsoids they span in 3D space. The essential behavior in the functional simulation is mainly located at the binding site: a hinge bending motion of helix 4 (residues 80-90), combined with a downwards motion of the loop consisting of residues 179-187 is visible. These residues are important for binding of the cutinase to a lipid surface, binding of substrates and release of products, as has been shown by site directed mutagenesis. High mobility is also visible for the N- and C-terminal, helix 8 (residues 152-158), residues 172-174, the loop between helix 3 and strand 2 (residues 64-66) and residue 33.
The essential motions of S-unf-sim are less localized and much larger (figure 3.7). The functional motions in the binding region are also observed in this simulation. Residues 179-187 show an additional motion in the binding site. This extra mobility suggests that a correlated opening or closing of the two binding loops is possible, required for function of the enzyme. The motion of helix 8 in S-unf-sim is much larger than and perpendicular to that in funct-sim and extends over the adjacent loop and helix 9 (residues 151-166). In contrast with the behavior in funct-sim, where only a rearrangement takes place, in S-unf-sim this loop detaches from the protein core. Helix 7 and the loop between helix 6 and 7 (residues 133-141) are moving in the direction of this large unfolding loop, probably a secondary effect due to the unfolding of this loop. Helices 12 and 13 (residues 191-213) are moving away from helix 3 (residues 52-66), which itself undergoes some motion. This correlated motion of the helices leads to a deformation of the protein. Other large, more local motions are occurring in the N- and C-terminal, loop 64-66 and residue 33.

3.3.3. Atomic distance monitoring

The rearrangements were also followed by monitoring the atomic distances between judiciously chosen atoms in course of the MD simulation. This analysis has no bias due to translational-rotational fitting, as opposed to RMSF and ED analyses. The motion of the binding site was monitored from two atoms in the middle of the two binding loops: the Cα-atom of Ala85 in helix 4 (the first binding loop) and the Cα-atom of Val184 in the loop between helix 10 and 11 (the second binding loop) for all simulations. This distance to the Cα-atom of Gly117, which is situated in the middle of strand 3 in the core of the sheet, is nearly constant for the atom in the first loop, which is in agreement with the axial rotation of helix 4, and is decreasing for the atom in the second binding loop, which is in agreement with the closing of the second binding loop (data not shown).

The atomic distances for the Cα-atoms of residues 151-166 in the large loop consisting of helix 8 and helix 9 with respect to the core (Cα-atom of Gly117) of the enzyme, all increase, some even up to 8 Å for S-unf-sim, which indicates that the loop indeed detaches from the core of the protein. This is in contrast to the funct-sim, in which only a small increase or decrease in these distances occur, corresponding to a conformational rearrangement in the loop.

Only in S-unf-sim the distances between the Cα-atoms of residues 132-134 (loop connecting helix 6 and 7) and the Cα-atom of Gln103 (middle of helix 5) are increasing, indicating a conformational change of this loop during unfolding.

The separation between helix 3 and helix 13 during unfolding is derived from the increase in distances between the Cα-atoms of residues 201-206 in helix 13 and those of residues 59 and 60 in helix 3 in S-unf-sim. In funct-sim these distances are constant.
Figure 3.7: The essential eigenvectors of funct-sim and S-unf-sim presented as thermal ellipsoids. Size and anisotropy of Cα movements are presented. The residues participating in the functional behavior are colored red, the residues involved in the conformational adaptation are colored green, the ones involved in the initial unfolding are colored magenta and the rest is colored yellow.
3.3.4. Hydrophobic patches

Large hydrophobic patches, as detected using QUILT, over the whole funct-sim are observed for the binding loops (ranges 81-88 and 173-190) including the top of ox-
yanion hole residue 42 and helix 8 (153-156) and are represented on the Connolly sur-
face in figure 3.8a. Smaller patches are identified around loops at the bottom of the
enzyme (29-30, 64-67, 111) and along helix 12.

Figure 3.8: Front and back view of the solid solvent accessible surface (grey): For the crystal
structure with the surface of the atoms participating in the four largest patches colored black
(a). For the 600 ps S-unf-structure (b) and the crystal structure (c) with the surface of the at-
oms participating in the three largest patches of the unfolding structure colored black.
Chapter 3

The large hydrophobic patches found in funct-sim, are found in S-unf-sim as well and have approximately the same size (not shown). Patches, three to four times larger than the largest patch of funct-sim, are developing during this simulation between helices 12 and 13, and helix 3 (figure 3.8b-front), as well as between helix 5 and helix 7 (figure 3.8b-back). Especially the first patch is growing fast and is becoming very large in the course of the unfolding simulation. Both patches are growing from the bottom to the top of the enzyme. In funct-sim much smaller patches were recognized in the corresponding regions (figure 3.8c).

Comparison of the Connolly surfaces\(^27\) of different frames of S-unf-sim shows the formation of two water-filled crevices. These are localized in the same regions as the largest hydrophobic patches for this simulation. After 600 ps the crevices lay parallel over the whole length of the β-sheet moving their facing helices apart. The largest crevice is formed between helix 12 and 13, and helix 3. The other crevice is situated between the last turn of helix 6 (residues 126-132) and helix 5 (residues 91-108). The formation of this crevice is probably due to loss of packing at the other side of helix 6 and 7 as result of the unfolding of residues 151-166.

3.4. DISCUSSION

3.4.1. Functional behavior

The opening of the active site is of biological relevance as only in this way substrates can enter and products can leave the binding site. Funct-sim shows large motions involving both binding loops. The first loop shows an axial rotation of helix 4 and adjacent residues, with residues 80 and 90 acting as hinge points. The second binding loop, residues 179-187, undergoes a rigid body motion, which is correlated with the partial unfolding of the top of helix 12. These large motions are however not obvious from the crystallographic B-factors, as the first binding loop has relatively low B-factors. The RMSF from the simulation for residues 42-45 is also relatively large compared with crystallographic B-factors. This is of biological relevance, as Ser42 forms the oxyanion hole together with Gln121.\(^28\) B-factors in the first binding loop and for residues 42-45 are probably low due to crystal contacts. In contrast, the mobility in these regions is confirmed by NMR, both for the residues in the two binding loops and for the residues forming the oxyanion hole.\(^29,30\)

In both unfolding simulations similar motions as in funct-sim are visible for the binding loops and residues 42-45, but with a larger amplitude (not shown). In S-unf-sim the axial rotation of helix 4 is larger than in T-unf-sim, increasing the exposed hydrophobic surface. The different environment in this simulation is probably the driving force for the exposure of more hydrophobic surface. In T-unf-sim the higher amplitude is caused by the higher temperature, which increases the flexibility overall.
3.4.2. Unfolding behavior

In all unfolding simulations we performed so far, the same parts of the enzyme unfold. For the two simulations analyzed in detail in this article (S-unf-sim and T-unf-sim) the RMSD is comparable, which is an indication that the enzyme unfolds to approximately the same extent. The evaluation of secondary structure, radius of gyration and solvent accessible surface area during the simulation, however, clearly shows that the order and the direction of unfolding motion may be different. In S-unf-sim the solvent is less hydrophilic, which results in a less unfavorable interaction with the hydrophobic atoms of the protein. Hence the interaction of the solvent with hydrophobic side chains of the enzyme competes more with hydrophobic enzyme-enzyme interactions, breaking some weaker ones. In T-unf-sim fluctuations are larger and now the gain in entropy pays for breaking weak interactions between secondary structures: high entropy states are relatively favored. In the latter simulation this gives rise to unfolding while keeping the hydrophobic side chains together within the secondary structure elements, whereas in the former this is accompanied by a separation of the secondary structure elements. Although the behavior is different, the same regions respond to the unfolding conditions, indicating the existence of intrinsic weak spots in the structure of cutinase. It is thought that the apolar tails of anionics bind to concave hydrophobic patches at the surface of the protein. Therefore it is expected that anionics will also induce conformational changes that are characterized by exposure of hydrophobic surfaces. As S-unf-sim mimics this behavior better than T-unf-sim, the former was analyzed in more detail to understand the unfolding by anionics.

As expected, S-unf-sim encompasses the functional motions of the regions covering the active site similar to what is seen in funct-sim (figure 3.7, see discussion above). More strikingly, large correlated motions are recognized in S-unf-sim which are not present in funct-sim, resulting in the initial unfolding of the protein. Unfortunately, it was not possible to assign single ED eigenvectors exclusively to either type of behavior.

Three different regions are found to be important for the unfolding of cutinase. First, an important region for the initial unfolding is the growth of a hydrophobic crevice between on the one hand helices 12 and 13, and on the other hand helix 3 and the following loop (figure 3.8b-front), which starts from a hydrophobic patch which was already present at the bottom of the enzyme in the starting structure (see figure 3.8c-front). This crevice grows by correlated motions of helices 12 and 13, away from helix 3 (see figure 3.7). In the middle of helix 3 an H-bond is lost between the main chain of residues 55 and 51 (not shown), producing a kink that moves the C-terminal part and the following loop away. The atomic distance monitoring (not shown) and the anisotropy of the essential motions (figure 3.7) show that this kink motion does not take place in funct-sim. Experimental evidence that this region is indeed a weak spot in the cutinase structure is obtained from aspecific proteolysis in the presence of anionics, yielding, amongst others, cuts at residues 56 and 205 in this unfolding region (M. R. Egmond, personal communication). The homologous cutinase of Aspergillus oryzae has an additional cysteine-bridge which connects the part after the kink in helix 3 to...
strand 2 (see figure 3.9, for the position of this bridge), which might stabilize this region.

Second, the largest conformational change in the ED analysis of S-unf-sim is seen for
residues 151-166, the large loop containing helices 8 and 9 (figure 3.7). From the atomic distance changes in both simulations it is shown that these residues move away from the core of the protein in S-unf-sim, whereas in funct-sim only a small conformational change takes place in this region. The anisotropy in the essential dynamics analyses for these residues is clearly different between the simulations. The protruding loop formed in S-unf-sim, creates no large hydrophobic patch (figure 3.8b-back), because the loop and its interface to the core of the protein consist largely of hydrophilic residues. Therefore this weak spot in the molecular architecture is expected to be less sensitive to unfolding by detergents.

Third, because of the conformational change that takes place for helix 7 and the turn between helix 6 and 7 (figure 3.5 and figure 3.7), an extensive hydrophobic crevice grows from the bottom of the enzyme between helix 5 and the last turn of helix 6 (residues 126-132) (figure 3.8b-back). However, the secondary structure of helix 7 is maintained: the loop between helix 6 and 7 allows the helix to move as a rigid body. This conformational change may be triggered by the conformational change of residues 151-166.

Although unfolding due to binding of anionic surfactants to the hydrophobic active site can not be rigorously excluded, this type of unfolding is not supported by our simulations as no large secondary or tertiary contacts are lost in this region and no real large hydrophobic patch is formed during unfolding.

As the two main hydrophobic crevices, which are found in S-unf-sim are large enough for water to enter, the hydrophobic tails of anionic surfactants fit as well. The small hydrophobic patches found in the same regions in funct-sim may bind these tails and increase the hydrophobic surface area as in S-unf-sim and unfold the molecule in this way. As anionics have a negative head group, the presence of a positive charge in the extension of a hydrophobic crevice forms an extra element of recognition. In the first unfolding region, Arg196 is found in the extension of the hydrophobic crevice. This strongly suggest that the first region is the most important weak region for unfolding by anionics.

### 3.5. CONCLUSIONS

Essential dynamic analysis is a powerful tool to understand the concerted motions in a molecular dynamics trajectory. We showed that it is possible to distinguish between the functional and initial unfolding motions of cutinase, and to relate these to distinct regions in the molecular structure. The regions are summarized below and presented in figure 3.9.

The essential dynamics analyses of the functional and the unfolding simulations show large correlated motions of the loops consisting of residues 80-90 (containing helix 4) and of residues 179-187, previously identified as binding loops, that could facilitate the entering of substrate and leaving of product from the binding site (figure 3.7). Also residues 42-45, with Ser42 being part of the oxyanion hole, are found to take part in
Chapter 3

the functional motion.

Three main regions are recognized which are likely to be involved in the initiation of unfolding in cutinase.

1. Helices 12 and 13, and helix 3 move apart in the course of unfolding, giving rise to a large hydrophobic crevice. A kink motion of in the middle of helix 3 (between residues 51-55) and a conformational change in the loop following helix 3 (residues 64-66) characterize this motion.

2. The loop connecting strands 4 and 5 of the parallel β-sheet (residues 151-166) moves away from the core of the enzyme. This region may be of less importance for unfolding by anionic surfactants as it does not give rise to an increase in hydrophobicity.

3. The loop connecting helix 6 and 7 and helix 7 move away from helix 5, probably triggered by the unfolding of residues 151-166. This gives rise to a large hydrophobic crevice as well.

As it is possible to assign regions which are only sensitive to unfolding but have no influence on the function of the enzyme, it should be possible to design variants to improve the stability without affecting functionality. These variants should decrease the sensitivity to unfolding by anionic surfactants, without hampering the function. An obvious proposal for a stabilizing mutant of \textit{Fusarium solani pisi} cutinase is the introduction of the additional cysteine-bridge as found in \textit{Aspergillus oryzae} cutinase or mutation of Arg196 (see discussion above). Work on these and other variants is in progress.

Acknowledgments

We thank A. B. M. Linssen for his help with the ED analysis, P. Hünenerberger for his help with funct-sim, and H. Peters for the model of \textit{Aspergillus oryzae} cutinase. We thank J. J. Prompers and M. R. Egmond for discussions and suggestions.

REFERENCES


Functional and unfolding motions of cutinase

Chapter 3


